Pattern Recognition Letters 14 (1993) 155-161 February 1993

# A coal micrograph image understanding system

D.P. Mukherjee, D.K. Banerjee, B. Uma Shankar and D. Dutta Majumder

*National Centre for Knowledge Based Computing, Indian Statistical Institute, Calcutta, India*

Received 24 January 1992
Revised 1 June 1992

*Abstract*

Mukherjee, D.P., D.K. Banerjee, B. Uma Shankar and D. Dutta Majumder, A coal micrograph image understanding system, Pattern Recognition Letters 14 (1993) 155-161.

A pattern recognition based technique has been used to classify the different constituent macerals of coal. The method has two modules, an off-line training module and an on-line classification module. The result shows that a simple pattern recognition technique with a bare minimum hardware is a faster, cost efficient, highly productive coal inspection procedure contrary to the subjective and tedious current inspection practice.

*Keywords.* Rank, macerals, classification, minimum distance classifier.

## 1. Introduction

Industrial use of coal depends on its quality. There are various standards for the quality classification of coal into various *ranks*[1] Basically, two parameters, *fixed carbon content* and *calorific value*, are used for this classification. These systems serve well for the general utilization of coal by the metallurgical coke industries (e.g., Steel), who are interested in the fixed carbon content and coking ability of coal, and by power generation industries, to whom the heating ability is of primary concern. Many existing methods of classification are based on the reflectance of coal.

Coal is an aggregate of heterogeneous substances composed of organic and inorganic materials. Optically homogeneous, microscopically recognizable organic constituents of coal are called *macerals*. There are three major groups of macerals, namely, vitrinite, exinite (liptinite) and inertinite. For most purposes, the reflectance of vitrinite is used as a measure of the reflectance of coal (Davis (1978)).

The current practice (Ting (1978)) of coal sample analysis with the help of micrographs is a very time-consuming process because 1000 points per sample (from 100 micrographs corresponding to different views) need to be taken. Each micrograph is manually inspected by an analyst and it takes almost a full shift for an operator to inspect one sample. Such operators are scarce and the overall operation is repetitive, strenuous, subjective and often not very accurate. Moreover, the number of samples per day are much too high compared to the number of trained operators and the microscope systems available. Keeping this in view, the

*Correspondence to*: Dr. D.P. Mukherjee, ECSU/KBSC, Indian Statistical Institute, 203 B.T. Road, Calcutta 700 035, India.

[1] The word *rank* is generally used to represent the degree of metamorphism or progressive alteration, in the natural series from lignite to anthracite.

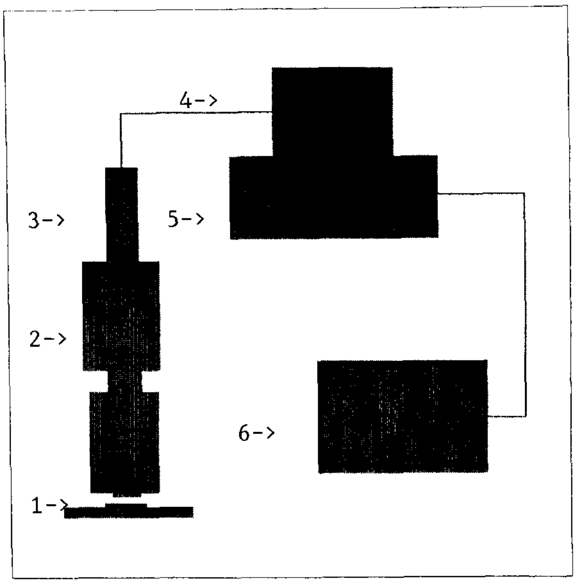


Figure 1. Inspection system components. 1. Micrograph sample, 2. Microscope, 3. CCD camera, 4. PC and digitizer, 5. Algorithm and knowledgebase, 6. Display terminal.

following objectives were laid down for the project: (i) to standardize the analysis on an objective basis; (ii) to automate the analysis of the samples to a high degree.

The proposed solution to the problem was sought with the help of digitized images of the coal

Figure 2. Micrograph plate I (red band).

micrographs. The classification of the different constituent parts has been done by using a simple clustering algorithm, the minimum distance classifier and, once the different components have been reliably separated, the counting and computation of relative proportions are straightforward. The main contribution of the paper is the evidence of an important industrial application. It

Table 1
Output from the training module

| | Cluster means | | |
|---|---|---|---|
| Clusters | Red | Green | Blue |
| Background | 56 | 45 | 51 |
| Inertinite | 177 | 142 | 100 |
| Exinite | 96 | 75 | 64 |
| Vitrinite | 134 | 109 | 82 |

| | Cluster standard deviations | | |
|---|---|---|---|
| Clusters | Red | Green | Blue |
| Background | 10.0 | 7.0 | 7.2 |
| Inertinite | 17.7 | 14.0 | 11.4 |
| Exinite | 12.4 | 10.7 | 9.7 |
| Vitrinite | 12.1 | 9.8 | 9.0 |

| | Distance between vitrinite and other clusters | | |
|---|---|---|---|
| Clusters | Background | Inertinite | Exinite |
| Vitrinite | 106.0 | 56.7 | 54.7 |

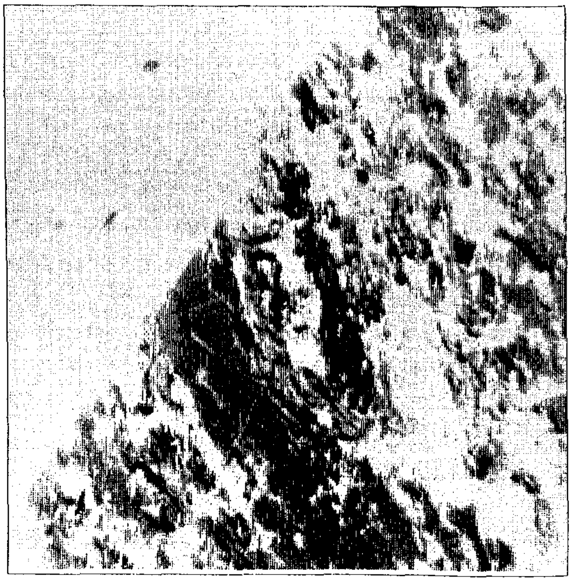

Figure 3. Micrograph plate I (green band).

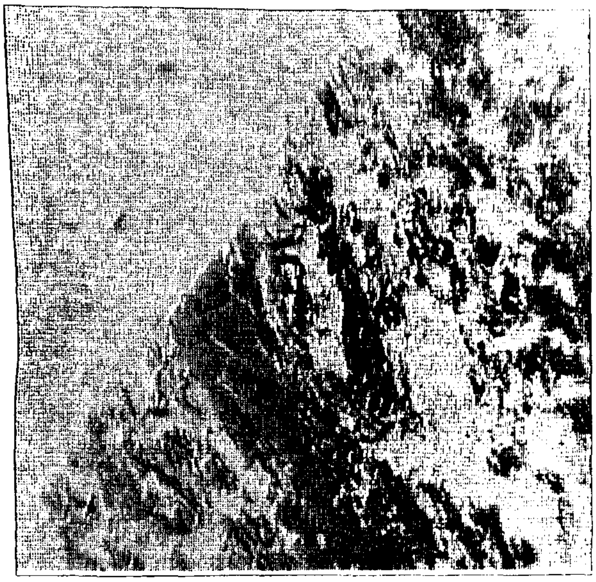

Figure 4. Micrograph plate I (blue band).

amply demonstrates the power of even a very simple P.R. technique in industrial problems.

Organization of the paper is as follows. We have defined the problem in the next section, followed by the proposed solution methodology. The results obtained by applying the P.R. technique are described subsequently, followed by discussions and concluding remarks.

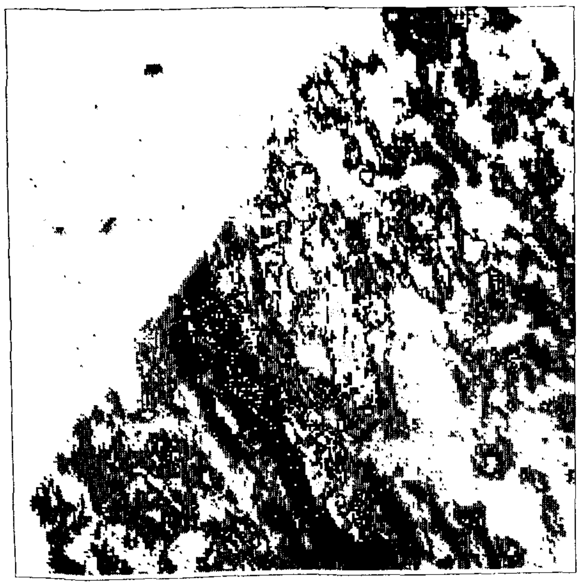

Figure 5. Plate I classified image.

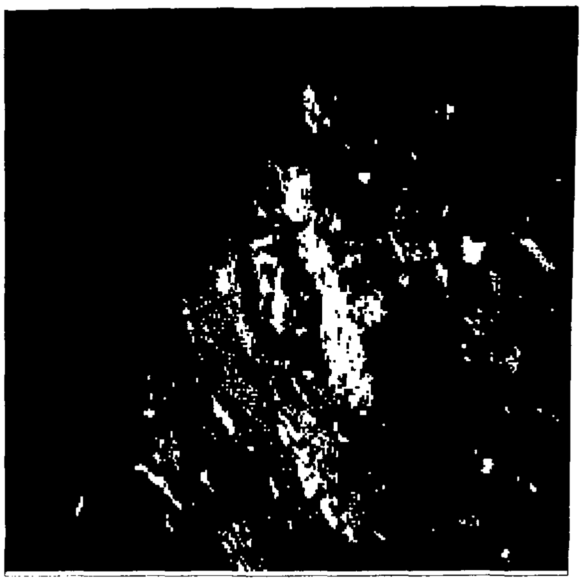

Figure 6. Inertinite in plate I.

## 2. Problem definition

The two basic problems as given by the user industry involved in coal sample analysis are:

(a) to determine the quantities of reactive and inert substances in the coal by point-counting, that is to find the relative percentages of vitrinite, exinite and inertinite;

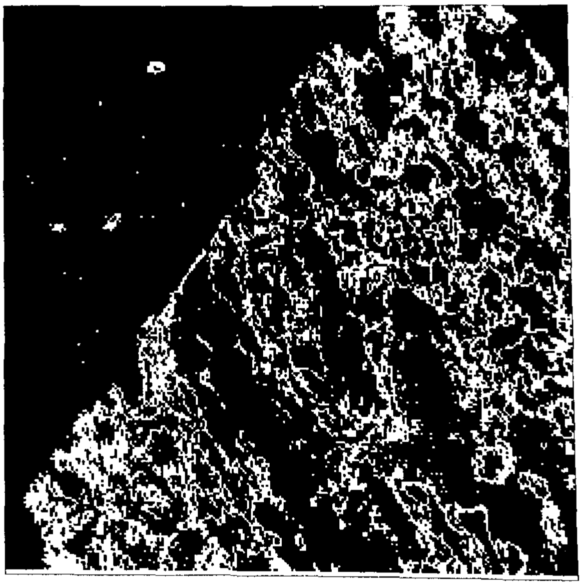

Figure 7. Exinite in plate I.

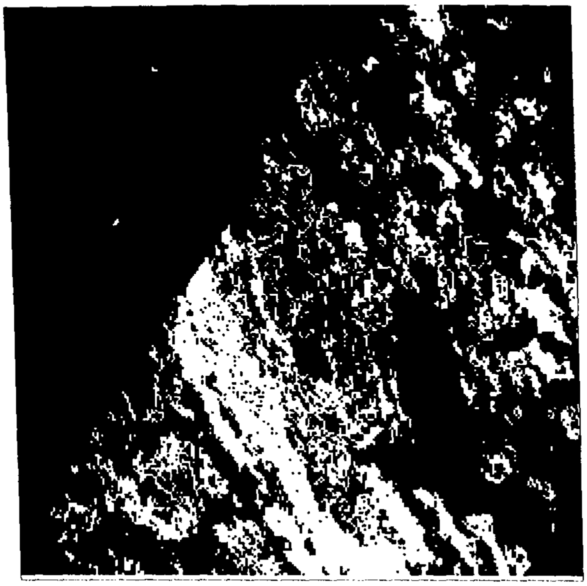

Figure 8. Vitrinite in plate I.

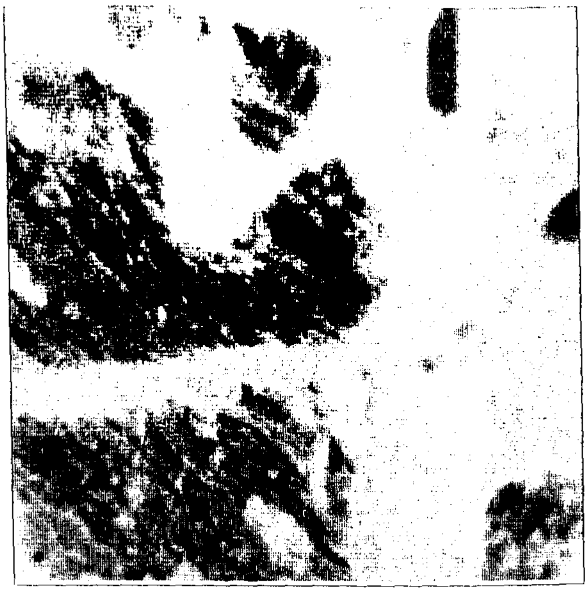

Figure 10. Micrograph plate II (green band).

(b) to determine the maturity of the reactive substances by reflectance of light, that is, primarily to find the mean reflectance of vitrinite.

## 3. Proposed solution methodology

The basic task here is visual inspection where speed and accuracy are essential. A human inspector looks at a micrograph and sees various coal constituents present in the field of view. The fields of view are usually distributed in different constituent patches of different size and the inspector makes a guesstimate of the relative proportions of these areas present in the micrographs. The automatic approach proposed here will first capture and digitize the microscopic image and store it in the computer memory. Then, using pattern recog-

Figure 9. Micrograph plate II (red band).

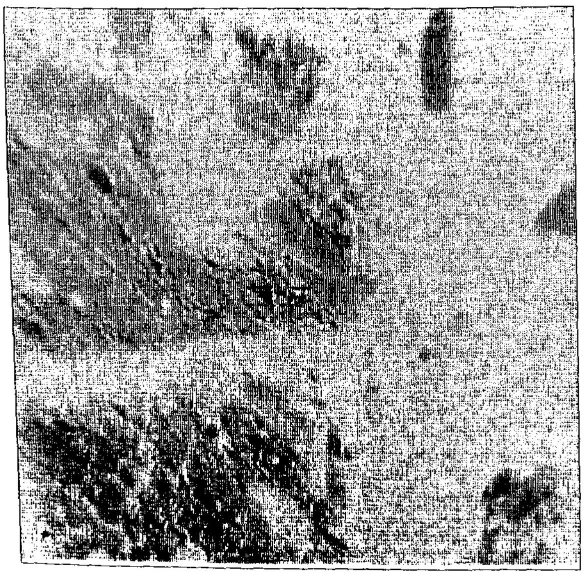

Figure 11. Micrograph plate II (blue band).

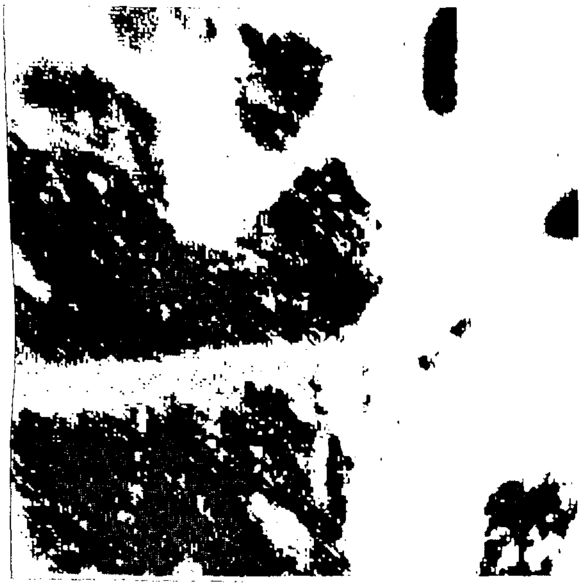

Figure 12. Plate II classified image.

nition techniques, it should detect different patches corresponding to different constituents in the digitized image from some a priori knowledge stored in the system. Finally, it will accurately compute the areas of the patches and the relative proportions of different constituents and grade the coal accordingly.

*Hardware*

A Leitz-Orthoplan Microscope has been used at a magnification of 500X. To facilitate computer-based analysis of the micrographs, the pictures were taken using a PC-AT with a Matrox card mounted on it, with the help of a Pullnix camera which gave digitized data in three spectral bands namely red, green and blue. The different system components have been pictorially depicted in Figure 1.

*Training module*

There are four classes (three for the constituent parts and one for the background) in the coal sample and we do not have any other information to set any classification criteria. Under these circumstances, we use an unsupervised clustering algorithm to find the necessary parameters which are useful in a future clustering process. Many clustering criteria have been studied. However, all of them reflect the assignment of points to clusters in such a way that the distance between points within a cluster are minimum and the distance between clusters are maximum. With this background, we proceed with the following clustering algorithm (Duda and Hart (1978)):

1. Select $c$ vectors (typically 4) to serve as initial cluster centres. The selection is arbitrary, except that no two of the initial cluster centres are sufficiently close to each other. The number of clusters $c$ must be specified by the analyst.
2. Assign each vector of data (red, green and blue graylevel values) to the nearest cluster centre. This step requires the specification of a suitable point-to-point distance measure. For this purpose a weighted Euclidean distance is used where the weights are inversely proportional to the standard deviations within the clusters.
3. Once the entire data set has been assigned to clusters, compute the new cluster centres and the new cluster standard deviations.
4. If the new cluster centres are sufficiently close to the old cluster centres then stop. Otherwise set old cluster centres equal to the new cluster centres and repeat from step 2.

Figure 13. Inertinite in plate II.

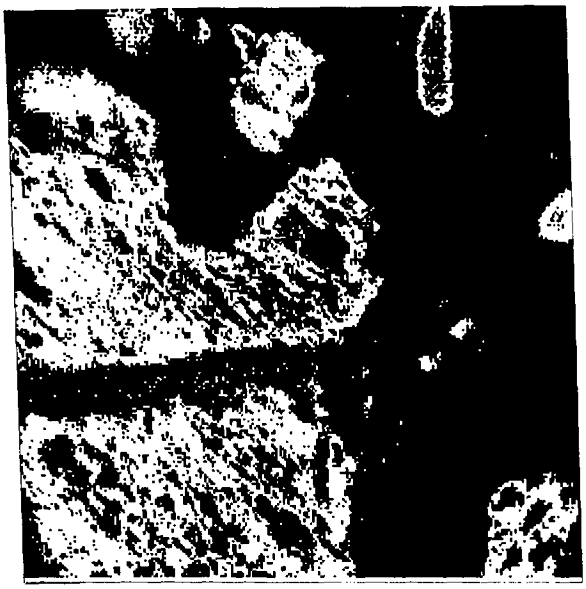

Figure 14. Exinite in plate II.

We tried the above algorithm on four images and found that the cluster means and variances are not significantly different from one image to the other. This suggests that there are four points in the spectral space representing the four clusters, which could be used for faster classification of the digitized image of the micrograph.

*Classification module*

Using the means and standard deviations learned above, the classification module assigns each pixel to the nearest cluster on the basis of the weighted Euclidean distance between the pixel and the cluster centre.

Finally, when all the pixels have been classified they are counted and the relative proportions of the cluster size are computed. The mean and standard deviation of the reflectance of vitrinite are also computed which are useful for determining the maturity of the coal sample.

## 4. Results and discussion

We used four micrograph plates each of $256 \times 256$ pixels. The output of the training module is shown in Table 1 where the stabilized cluster means and standard deviations for all the bands are presented. Table 1 also gives the inter-cluster distances between vitrinite and other clusters. Figures 2–4 are the three bands (red, green and blue) of plate I. Figure 5 is the classified image based on them. Figures 6–8 are the individual clusters, namely, inertinite, exinite and vitrinite, respectively. Similarly, Figures 9–15 are the three band images as well as the classified images of plate II respectively. Due to space constraints, figures for only plates I and II are included. Table 2 presents the final results where constituent percentages and the mean and standard deviation of the vitrinite cluster are given, which is useful in ranking the coal. From Table 2 we find that the means and standard deviations of vitrinite in all the plates are fairly close to each other. They are also close to the cluster mean for the vitrinite cluster shown in Table 1.

The training module is run off-line and so it is not considered in the final analysis of time. In the classification stage, which is on-line, the processing time per plate is 1 minute and 25 seconds (approx.) including I/O operations. Compared to the current practice it is clear that our image understanding system will immensely increase the productivity of the industry. It is to be noted that the current algorithm is implemented on the MicroVAX II on Ultrix platform.

Figure 15. Vitrinite in plate II.

Table 2
Coal micrograph plate statistics

| Plates | Percentages of constituents | | | |
|---|---|---|---|---|
| | Inertinite | Exinite | Vitrinite | Background |
| Plate I | 5 | 19 | 20 | 46 |
| Plate II | 0 | 29 | 10 | 61 |
| Plate III | 5 | 20 | 20 | 55 |
| Plate IV | 3 | 44 | 33 | 20 |

| Plates | Vitrinite means in the four plates | | |
|---|---|---|---|
| | Red | Green | Blue |
| Plate I | 133 | 108 | 82 |
| Plate II | 126 | 101 | 79 |
| Plate III | 132 | 107 | 81 |
| Plate IV | 133 | 105 | 76 |

| Plates | Standard deviations of vitrinite clusters | | |
|---|---|---|---|
| | Red | Green | Blue |
| Plate I | 12.1 | 9.8 | 9.0 |
| Plate II | 10.6 | 8.5 | 8.2 |
| Plate III | 12.1 | 9.9 | 9.4 |
| Plate IV | 12.2 | 9.6 | 8.7 |

## 5. Conclusion

A pattern recognition based technique has been used to classify the different constituent macerals of coal. The method has two modules, an off-line training module and an on-line classification module. The speed of the system is considerably faster than the current technique of analysis and has an objective basis. Moreover, the data of three bands of the same micrograph plate are highly correlated and it might be possible to work with a single band image. This will further reduce the processing time.

It is possible to increase the number of clusters allowing for other macerals of coal or to find subclusters within a particular cluster.

There is scope for improvement by induction of domain specific expertise, especially for fine-tuning the different parameters of clustering. Lighting conditions may be an important parameter against the system performance though the chances are less as the microscope is used in a controlled-lighting environment.

## Acknowledgments

This project was supported by DoE, Govt. of India and UNDP under the grant IND/85/072. The authors are indebted to Dr. S.K. Parui and Dr. Amita Pal and the Tata Iron and Steel Company Ltd.

## References

Davis, A. (1978). The reflectance of coal. In: C. Karr, Jr., Ed., *Analytic Methods for Coal and Coal Products, Vol. I.* Academic Press, New York.

Duda, R.O. and P.E. Hart (1978). *Pattern Classification and Scene Analysis.* Wiley, New York.

Ting, F.T.C. (1978). Petrographic techniques in coal analysis. In: C. Karr, Jr., Ed., *Analytic Methods for Coal and Coal Products, Vol. I.* Academic Press, New York.